

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-219590

(43)公開日 平成7年(1995)8月18日

(51) Int.Cl.⁶

G10L 3/00

識別記号

571 H

531 D

片内整理番号

FI

技術表示箇所

G O 6 F 17/30

9194-5L.

G O B F 15/ 40

370 E

審査請求 未請求 請求項の数10 O.L (全 9 頁)

(21) 出願番号 特願平6-9559

(22)出願日 平成6年(1994)1月31日

(71)出題人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72)發明者 酒井 桂一

東京都大田区下丸子3丁目30番2号キヤノン株式会社内

(72)発明者 池田 裕治

東京都大田区下丸子3丁目30番2号キヤノン株式会社内

(72) 發明者 藤田 稔

東京都大田区下丸子3丁目30番2号キヤノ
ン株式会社内

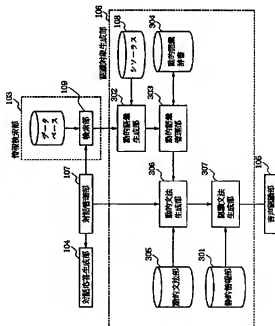
(74)代理人 井理士 丸島 儀一

(54)【発明の名称】 音声情報検索装置及び方法

(57) 【要約】

【目的】 対話の進行に応じて、音声認識のための語彙を動的に変更し、要求に応じて以前に使用した語彙に戻すことが可能な音声情報検索装置を提供する。

【構成】 音声情報検索装置に、音声認識の対象となる語彙を記録し、その記憶内容参照して入力音声情報を認識する音声認識部105と、その認識結果を解釈する対話管理部107と、その解釈結果に基づいて情報を検索する情報検索部103と、その検索結果に基づいて応答を生成する対話応答生成部104と、対話管理部107による解釈結果及び情報検索部103による検索結果に基づいて、動的対話管理部303を介して動的語彙辞書304を参照して、音声認識部105に記録された音声認識の対象となる語彙と、動的対話管理部303に管理された使用済みの語彙の履歴とを更新するための動的語彙生成部302とを具備する。



1

【特許請求の範囲】

【請求項1】 音声情報を入力する入力手段と、
音声認識の対象となる語彙を記憶する対象語彙記憶手段と、

と、
以前に音声認識の対象となった語彙の履歴を記憶管理する語彙履歴記憶手段と、

前記対象語彙記憶手段を参照して前記入力手段より入力された音声情報の認識を実行する認識手段と、

該認識手段による認識結果を解釈する解釈手段と、

該解釈手段の解釈結果に基づいて、情報を検索する検索手段と、

該検索手段による検索結果に基づいて、応答を生成する生成手段と、

該生成手段により生成された応答を出力する出力手段と、

前記解釈手段による解釈結果及び前記検索手段による検索結果に基づいて、前記語彙履歴記憶手段を参照して、前記記憶手段に記憶された音声認識の対象となる語彙と、前記語彙履歴記憶手段に記憶された語彙の履歴とを更新する更新手段とを具えたことを特徴とする音声情報検索装置。

【請求項2】 前記更新手段は、前記解釈結果が、音声認識の対象となる語彙を以前の語彙へ変更する要求であるとき、前記語彙履歴記憶手段を参照して、前記対象語彙記憶手段の語彙を以前の語彙へ変更することを特徴とする請求項1記載の音声情報検索装置。

【請求項3】 前記更新手段は、前記解釈結果が、音声認識の対象となる語彙を直前の語彙へ変更する要求であるとき、前記語彙履歴記憶手段を参照して、前記対象語彙記憶手段の語彙を直前の語彙へ変更することを特徴とする請求項1記載の音声情報検索装置。

【請求項4】 前記対象語彙記憶手段は、前記更新手段の更新対象となる動的な語彙と、前記更新手段の更新対象外となる静的な語彙とを記憶することを特徴とする請求項1記載の音声情報検索装置。

【請求項5】 前記更新手段は、語彙を当該語彙間の階層関係とともに記憶する辞書を有し、当該語彙及び階層関係を参照して、前記対象語彙記憶手段に記憶された語彙を更新することを特徴とする請求項1記載の音声情報検索装置。

【請求項6】 音声情報を入力する入力工程と、
音声認識の対象となる語彙を記憶する対象語彙記憶部を参照して、前記入力工程において入力された音声情報の認識を実行する認識工程と、

該認識工程による認識結果を解釈する解釈工程と、
該解釈工程の解釈結果に基づいて、情報を検索する検索工程と、

該検索工程による検索結果に基づいて応答を生成する生成工程と、

該生成工程により生成された応答を出力する出力工程

2

と、
前記解釈工程による解釈結果及び前記検索工程による検索結果に基づいて、以前に音声認識の対象となった語彙の履歴を記憶管理する語彙履歴記憶部を参照して、前記対象語彙記憶部に記憶された音声認識の対象となる語彙と、前記語彙履歴記憶部に記憶された語彙の履歴とを更新する更新工程とを具えたことを特徴とする音声情報検索方法。

【請求項7】 前記更新工程において、前記解釈結果が、音声認識の対象となる語彙を以前の語彙へ変更する要求であるとき、前記語彙履歴記憶部を参照して、前記対象語彙記憶部の語彙を以前の語彙へ変更することを特徴とする請求項6記載の音声情報検索方法。

【請求項8】 前記更新工程において、前記解釈結果が、音声認識の対象となる語彙を直前の語彙へ変更する要求であるとき、前記語彙履歴記憶部を参照して、前記対象語彙記憶部の語彙を直前の語彙へ変更することを特徴とする請求項6記載の音声情報検索方法。

【請求項9】 前記対象語彙記憶部に、前記更新工程の更新対象となる動的な語彙と、前記更新工程の更新対象外となる静的な語彙とを記憶させることを特徴とする請求項6記載の音声情報検索方法。

【請求項10】 前記更新工程では、語彙を当該語彙間の階層関係とともに記憶する辞書に記述された当該語彙及び階層関係を参照して、前記対象語彙記憶部に記憶された語彙を更新することを特徴とする請求項4記載の音声情報検索方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、入力された音声情報に基づいて、データベース上の情報を検索し、その結果を出力する装置に関するものである。

【0002】

【従来の技術】人間と人間の間で行われる情報交換の媒体の中で、最も自然に使われるのが音声である。一方、計算機の飛躍的な進歩により、計算機は数値計算のみならず、様々な情報を扱うようになってきている。そこで、人間と計算機の間でも、情報交換の媒体として音声を利用したいという要求がある。

【0003】このような要求に応えるべく、音声によって入力された情報に基づいて情報検索を行う音声情報検索装置が登場してきている。

【0004】

【発明が解決しようとしている課題】しかしながら、上記従来の音声情報検索装置には、予め登録したごく一部の単語や文を用いてしか、音声による検索ができず、更に、ある対話の状態においては、その状態に対して予め登録された対話内容しか認識できないため、自然な対話を行うことができないという問題点も生じていた。このため、データベース上のあらゆる検索項目を自然に検索

50

3

できないという問題が生じていた。

【0005】また、一般に、対話により自然に検索を行う時には、対話のどの時点でも入力されるユーザ入力文が存在する。例えば、旅に関する情報検索における対話では、「九州地方にあるゴルフ場を知りたい。」等の非常にグローバルな質問や、「温泉にはどんな項目がありますか?」等の検索するための情報を得るための質問(メタ質問)がある。一方、対話が進むにつれて、詳細な内容に関わる質問、例えば、「箱根湯本温泉の電話番号を知りたい。」とか「群馬県吉井町のゴルフ上の立地条件は?」といった質問が入力されるようになる。

【0006】このように、対話のどの時点でも入力されるユーザ入力文を受け付けて音声認識するための音声認識情報と、対話が進むにつれて変わっていく入力を受け付けて音声認識するための音声認識情報とを同時に扱うためには、認識装置の巨大化や認識性能の低下、制御の複雑化が問題となっている。

【0007】そこで、本出願人は、上記問題を解決するために、先に出願した特願平5-117169号において、対話のどの時点でも現れる入力文を認識するための語彙(静的な語彙)と、対話の進行により変更される語彙(動的な語彙)とを区別して保持し、地名などの階層化された情報を有する辞書(シソーラス)の情報について、対話が進むにつれて、この階層構造に従い、順次、次の階層の情報を動的な語彙として用いるように、動的な語彙を変化させてゆくことで、認識性能を低下させたり、認識時間を増大させたりすることなく、対話全体を通じて多数の語句を認識可能とする音声情報検索装置及び方法を提案している。

【0008】しかしながら、上記先願の音声情報検索装置では、シソーラスの階層化された情報に基づいて、この階層を順に辿って動的な語彙を変更するために、ある時点の動的な語彙を、その語彙の階層とは離れた別の階層の語彙に変更するためには、それらの共通の祖に当たる階層から順次辿って行かなければならない。

【0009】例えば、地名に関するシソーラスを用いた場合には、その階層として、「地方名」、「都道府県名」、「市区町村名」があるが、「市区町村名」を別の「地方名」と誤認識した場合には、動的な語彙が切り替わってしまい、元の市区町村名を動的な語彙に戻すには、地方名から辿っていく必要がある。また、隣接している市区町村でも、その境界が地方の境界となっている場合には、その隣接した市区町村を認識語彙とするのに地方名から辿って行く必要がある。

【0010】このように、ユーザが意図しない動的な語彙に切り替わってしまった場合や、誤認識によって、ユーザが以前の対話で用いた動的な語彙を用いた場合に、動的な語彙を元に戻すために階層を順に辿っていくという、ユーザにとって冗長な対話処理を要するという問題があった。

4

【0011】

【課題を解決するための手段】上記課題を解決するために、本発明の音声情報検索装置は、音声認識の対象となる語彙を記憶する対象語彙記憶手段と、以前に音声認識の対象となった語彙の履歴を記憶管理する語彙履歴記憶手段と、前記対象語彙記憶手段を参照して前記入力手段より入力された音声情報の認識を実行する認識手段と、該認識手段による認識結果を解釈する解釈手段と、該解釈手段の解釈結果に基づいて、情報を検索する検索手段と、該検索手段による検索結果に基づいて、応答を生成する生成手段と、該生成手段により生成された応答を出力する出力手段と、前記解釈手段による解釈結果及び前記検索手段による検索結果に基づいて、前記語彙履歴記憶手段を参照して、前記語彙履歴記憶手段に記憶された語彙の履歴とを更新する更新手段とを具える。

【0012】また、本発明の他の態様によれば、音声情報検索方法に、音声情報を入力する入力工程と、音声認識の対象となる語彙を記憶する対象語彙記憶部を参照して、前記入力工程において入力された音声情報の認識を実行する認識工程と、該認識工程による認識結果を解釈する解釈工程と、該解釈工程の解釈結果に基づいて、情報を検索する検索工程と、該検索工程による検索結果に基づいて応答を生成する生成工程と、該生成工程により生成された応答を出力する出力工程と、前記解釈工程による解釈結果及び前記検索工程による検索結果に基づいて、以前に音声認識の対象となった語彙の履歴を記憶管理する語彙履歴記憶部を参照して、前記対象語彙記憶部に記憶された音声認識の対象となる語彙と、前記語彙履歴記憶部に記憶された語彙の履歴とを更新する更新工程とを具える。

【0013】

【作用】上記音声情報検索装置は、音声認識の対象となる語彙を記憶する対象語彙記憶手段を参照して、入力手段より入力された音声情報の認識を認識手段により実行し、その認識結果を解釈手段により解釈し、その解釈結果に基づいて、検索手段が情報を検索し、その検索結果に基づいて生成手段が生成した応答を出力手段が出力する。そして前記解釈手段による解釈結果及び前記検索手段による検索結果に基づいて、語彙履歴記憶手段を参照して、前記対象語彙記憶手段に記憶された音声認識の対象となる語彙と、前記語彙履歴記憶手段に記憶された語彙の履歴とを更新する。

【0014】また、上記音声情報検索方法は、音声認識の対象となる語彙を記憶する対象語彙記憶部を参照して、入力工程において入力された音声情報の認識を認識工程により実行し、その認識結果を解釈工程により解釈し、その解釈結果に基づいて検索工程により情報を検索し、その検索結果に基づいて生成工程により生成された応答を出力工程において出力する。そして、前記解釈工

5

程による解釈結果及び前記検索工程による検索結果に基づいて、前記語彙履歴記憶部を参照して、前記対象語彙記憶部に記憶された音声認識の対象となる語彙と、前記語彙履歴記憶部に記憶された語彙の履歴とを更新工程において更新する。

【0015】

【実施例】以下、図面を参照して本発明を詳細に説明する。

【0016】図1は、本発明の実施例の音声情報検索装置のブロック構成を示す図である。同図において、101は、マイクより入力された音声を入力／D変換して装置内部に取り込む一方、生成された応答をD/A変換してスピーカより出力する音声入出力部である。また、102は、検索結果、対話の流れ、次発話情報などを文字出力するための表示部である。103は、データベース110と検索部109を含む、そのデータベース110に格納されている情報を検索部109により検索して出力するための情報検索部である。104は、対話管理部107で管理される情報や情報検索部103で検索された検索情報より、出力形式の対話応答文を生成する対話応答生成部である。

【0017】また、105は、認識対象の情報である認識語彙（読みと表記の対）及び文法を参照して、音声入出力部101より入力された音声の認識を行う音声認識部である。106は、階層辞書であるシソーラス108を含み、それを参照して、次発話で用いられると予想される認識対象の情報を生成する認識対象生成部である。また、107は、音声認識部で認識された情報を解釈し、情報検索部103に検索指示を与え、次発話予測、対話応答を行う対話管理部である。

【0018】次に、以上のように構成される音声情報検索装置の動作を図2のフローチャートを参照して説明する。

【0019】まず、ステップS201では、音声入出力部101より音声入力を待ち、音声入力が行われると、この入力情報を音声認識部105に送る。ステップS202では、音声認識部105が、認識対象生成部106にて生成された情報を参照して、入力情報の音声認識を行う。

【0020】そして、ステップS203で、この認識の結果を対話管理部107に送り、対話管理部107が認識結果を解釈する。ステップS204にて、解釈結果が動的な語彙の変更要求であれば、ステップS209に進み、そうでなければ、ステップS205に進む。ステップS205では、検索条件が整っているかを判定し、整っていれば検索の指示が出され、ステップS206に進む。そうでなければ、不足情報を得るために、ステップS209へ進む。

【0021】ステップS206では、検索条件が整っているので、情報検索部103において、データベース1

6

10より情報を検索する。続いて、ステップS207で、対話管理部107や情報検索部103より出力される情報をもとに、動的な語彙の変更要求である場合の受理の応答、検索条件が整っていない場合の不足情報を得るための指示の応答、あるいは、検索を行なった場合の検索結果の応答を、対話応答生成部104が生成する。

【0022】ステップS208において、生成された応答を、音声入出力部101のスピーカより音声出力する、あるいは表示部102に文字出力する。一方、情報検索部103により検索された情報や対話管理部107にて管理される対話の状況をもとに、ステップS209では、次発話が発生すると予測される認識対象の情報（語彙及び文法）を、認識対象生成部106にて生成する。生成された認識対象の情報は、音声認識部105による次の発話の認識の際に参照できるように、音声認識部105内に保存される。認識対象の情報が生成されたら、ステップS201に戻り、再び、次発話の音声入力を待つ。

【0023】なお、このステップS209は、ステップS207～S208の対話応答の生成及び出力と並行して実行するようにしてもよく、そのようにすれば、待ち時間を短くすることができる。

【0024】また、図1の装置各部は、上述の如く動作する1つまたは複数の専用ハードウェアで構成してもよいし、その一部を、CPUが、メモリに記憶された図2のフローチャートに対応するプログラムを、メモリ上のワークエリアを利用しながら実行することで、実現するようにしてもよい。

【0025】図3は、認識対象生成部106の詳細構成を示す図である。図1と同じ部分には、同じ番号を付している。ただし、図1が共通バス形式による装置構成を示しているのに対し、図3は、装置各部間のデータの流れを明らかにするように構成されている。

【0026】前述したように、一般に、対話を自然に行う時には、対話のどこでも発声できる入力が存在する。例えば、旅に関する情報検索のための質問（メタ質問）が「九州地方にあるゴルフ場を知りたい。」等の非常にグローバルな質問や、「温泉にはどんな項目がありますか？」等の検索するための情報を得るための質問（メタ質問）がある。図3において、301は、このように対話のどこでも発声できる入力を受け付け、音声認識するための語彙及び文法が保持されている静的情報部である。

【0027】一方、対話が進むに連れて、詳細な内容に関わる質問、例えば、「箱根湯本温泉の電話番号を知りたい。」とか「群馬県吉井町のゴルフ場の立地条件は？」といった質問が入力されるようになる。

【0028】302は動的語彙生成部であり、このように対話が進むに連れて動的に変化する入力や音声認識するための語彙を、情報検索部103で検索された結果から、シソーラス108を参照して生成する。304は、

7

動的語彙を保持するための動的語彙辞書である。303は、動的語彙辞書304を管理する動的語彙管理部であり、動的語彙生成部302で生成された動的語彙は、動的語彙管理部303を介して動的語彙辞書304に保持される。

【0029】305は、動的な文法を保持する動的文法部である。306は、動的文法生成部であり、動的な語彙の変更要求、もしくは検索を行なった結果に応じて、動的語彙管理部303を介して動的語彙辞書304から得られる動的語彙、及び動的文法部305に保持されている文法から、語彙を含む動的な文法を生成する。

【0030】307は、認識文法生成部であり、動的文法生成部で生成された動的な文法に、静的情報部301に保持されている静的な情報を付け加えて、認識対象とする全ての情報を生成し、音声認識部105に送る。

【0031】以下、本実施例においては、シソーラス108に格納される階層化された単語情報を「地名」とした例について説明する。

【0032】「地名」は日本国内においても、関東地方、近畿地方などの「地方名」、東京都、大阪府などの「都道府県名」、堺市、千代田区、豊能町などの「市区町村名」といった階層関係を持っている。シソーラス108には、各「地方」に存在する「都道府県名」および各「都道府県」に存在する「市区町村名」の階層化した情報が格納されている。

【0033】図4に地名の認識語彙を動的に変更する処理のフローチャートを示す。

【0034】ステップS401で、静的な認識用の情報に、地名の最上位である「地方名」を保持し、ステップS402で、起動時の動的な認識用の情報に、「地方名」の次位に当たる「都道府県名」をすべて保持する。すなわち、起動時の時点で、音声認識可能な「地名」は、「地方名」と「都道府県名」である。

【0035】ステップS403では、音声認識を行ない、認識結果を対話管理部107が解釈し、検索条件を作成する。解釈結果が動的な語彙の変更要求であれば、ステップS405に移り、そうでなければ、ステップS406に移る。

【0036】ステップS405では、語彙の変更が要求されたので、動的語彙辞書304から動的語彙を取り出して、動的な認識用の情報に付け加えて、ステップS403に移る。

【0037】ステップS406では、作成された検索条件により情報検索部103で情報検索が行なわれ、ステップS407に移る。ステップS407では、検索条件中の地名が「地方名」であればステップS408に移り、「地方名」でなければステップS409に移る。ステップS408では、シソーラス108を参照して、検索結果が存在する「都道府県名」を動的な認識用の情報の地名語彙として、ステップS411に移る。

8

【0038】ステップS409では、検索条件中の地名が「都道府県名」であればステップS410に移り、「都道府県名」でなければ、地名語彙は変更せず、ステップS411に移る。ステップS410では、現在動的な認識用の情報として保持している「都道府県名」に加えて、シソーラス108を参照して、検索結果が存在する「市区町村名」を動的な認識用の情報の地名語彙として、ステップS411に移る。

【0039】ステップS411では、検索結果の件数がある閾値（ここでは、例えば10件とする）以下であれば、ステップS412に移り、この閾値より多ければ、ステップS413に進む。ステップS412では、検索の結果得られた具体的な名称を、動的な認識用の情報に付け加えてステップS413に移る。

【0040】本実施例に基づく対話例を図5、図7に示し、これらの対話例によって変化する音声認識語彙を図6、図8にそれぞれ示す。

【0041】図5中、usr3は、中国地方の都道府県名と、鳥取県の市区町村名を動的語彙とした検索中に、入力された「東伯町」が静的語彙の「東北地方」と誤認識された例である。このため東北地方の寺院が検索される。そして、sys3に示すように、動的語彙が東北地方の都道府県名となり、東伯町だけでなく、鳥取県までの動的語彙からはずれてしまう。このような場合、前述した先願の装置では、再び東伯町を動的語彙とするためには、「中国地方」、「鳥取県」を順に指定しなければならなかったが、本実施例では、sys2での動的語彙usr4の1対話により、sys4に示すように、再び東伯町を動的語彙とすることができる。

【0042】また、図7の対話の例では、図8に示すように、sys2での動的語彙は、関東地方の都道府県名と、群馬県の市区町村名であり、その後sys6では、甲信越地方の都道府県名と、長野県の市区町村名が動的語彙となっている。このような場合に、以前動的語彙であった「長野原町」を再び動的語彙とするためには、前述した先願の装置では、「関東地方」、「群馬県」を順に指定しなければならなかったが、本実施例では、usr7の1対話により、sys7での動的語彙をsys2の動的語彙に戻し、「長野原町」を再び動的語彙とすることができる。

【0043】以上のごとく、地名のように階層化した構造を持つ情報群を認識語彙とする場合、対話を通して着目する階層を変え、認識語彙を変更することで、認識性能を損なうことなく、音声対話を実現するとともに、誤認識や、対話の進行などにより、過去の動的語彙とは離れた階層の動的語彙に切り替わってしまった場合にも、過去の動的語彙に簡単に戻ることができるので、動的語彙の切替が効率的に行なえるようになった。

【0044】以上のように本実施例によれば、自然でしかも使い易い形で、音声入力による情報検索が実現でき

【0045】〔他の実施例〕上記実施例では、地名の階層として、「地方名」、「都道府県名」、「市区町村名」の3階層として説明したが、これに加えて、日本、アメリカ合衆国などの「国名」、更にはアジア、ヨーロッパなどの「地域名」など、階層を増やしても構わない。その場合、図4のステップS407～S410の分岐処理を増やすことによって対処できる。

【0046】また、上記実施例では階層化された情報として「地名」を用いたが、例えば、「会社組織の部署名」における「本部名」、「部名」、「課名」など、他の階層化された情報を利用しても、同様の処理が可能である。

【0047】また、以上の実施例では、階層化された情報を対象としたが、これに限らず、動的な情報として、単に複数のグループからなる情報の内のいくつかのグループを切り替え選択して利用する場合にも、使用したグループの履歴を管理することにより、上記実施例と同様に、過去に使用した動的な情報を、再度動的な情報とすることが簡単に実現できる。

【0048】

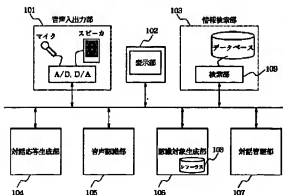
【発明の効果】以上説明したように、本発明によれば、音声認識のための認識情報を対話に応じて動的に変更できるとともに、認識情報を以前使用した認識情報へ簡単に戻せるようになり、自然で使い易い音声情報検索装置を実現できるという効果がある。

【図面の簡単な説明】

【図1】本発明による音声情報検索装置の実施例のブロック構成図である。

【図2】実施例の音声情報検索処理のフローチャートである。

【図1】



【図3】実施例の認識対象生成部の図である。

【図4】実施例の地名認識語彙の変更処理のフローチャートである。

【図5】実施例のユーザーとシステムの対話例を示す図である。

【図6】対話例によって変化する音声認識語彙の図である。

【図7】実施例のユーザーとシステムの対話例を示す図である。

【図8】対話例によって変化する音声認識語彙の図である。

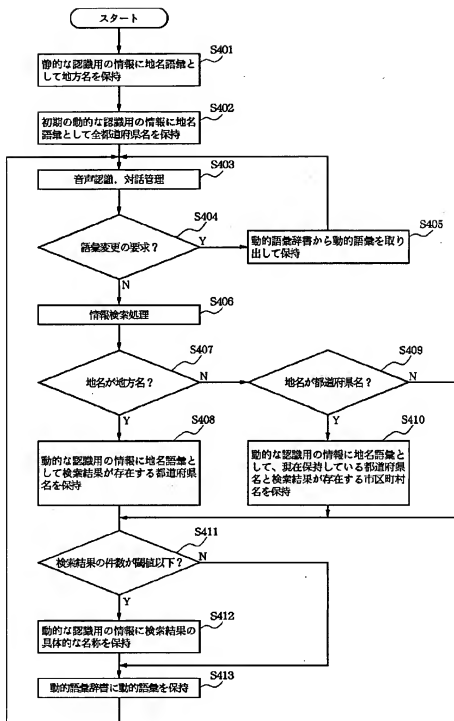
【符号の説明】

- 101 音声入力部
- 102 表示部
- 103 情報検索部
- 104 対話応答生成部
- 105 音声認識部
- 106 認識対象生成部
- 107 対話管理部
- 108 シソーラス
- 109 検索部
- 110 データベース
- 301 静的情報部
- 302 動的語彙生成部
- 303 動的語彙管理部
- 304 動的語彙辞書
- 305 動的文法部
- 306 動的文法生成部
- 307 認識文法生成部

【図5】

usr1:	中国地方にある今朝を知りたい。
sys1:	岡山県に104件、広島県に101件、山口県に186件、鳥取県に92件など、合計で629件あります。
usr2:	鳥取県では、
sys2:	鳥取県に19件、倉吉市に12件、鳥取県鳥取市に6件、若菜町に6件など、合計で87件あります。
usr3:	東北地方では、一宮市(トウハチマリ)を教えてください。
sys3:	宮城県に302件、青森県に282件、山形県に249件、秋田県に115件など、合計で1147件あります。
usr4:	記憶を元に戻せ。
sys4:	戻しました。
usr5:	山形町でね。
sys5:	山形市、山形県、山形市、山形県です。

【図4】



【图8】

[illegible]